

Hochwasserstatistik am BAFU – Diskussion eines neuen Methodensets

Eva Baumgartner, Marc-Olivier Boldi, Caroline Kan, Simon Schick

Zusammenfassung

Das Bundesamt für Umwelt (BAFU) hat sich entschieden, das bisher verwendete Standardverfahren des DVWK (1979) zur statistischen Abschätzung der Hochwasserwahrscheinlichkeit HQ_T abzulösen.

Am Fallbeispiel der Muota bei Ingenbohl wird ein Set von Auswertungsmethoden vorgestellt. Es beinhaltet ein semi-parametrisches Verfahren zur explorativen Datenanalyse sowie drei parametrische Verfahren zur Abschätzung der Hochwasserwahrscheinlichkeit HQ_T . Alle vorgestellten Methoden haben ihre Stärken und Schwächen. Da unbekannt ist, welcher Modellansatz die Realität am besten widerspiegelt, scheint es nicht zulässig, sich auf nur eine Methode zu beschränken. Das BAFU sieht daher vor, künftig für alle aktuellen Abflussmessstationen die Ergebnisse der hier vorgestellten Methoden in kurzen Berichten zusammenzustellen. Ziel des Artikels ist es, das angestrebte Methodenset der Fachwelt vorzustellen, eine Diskussion auszulösen und zahlreiche Rückmeldungen zu erhalten.

1. Einleitung

Das Wissen über Hochwasserwahrscheinlichkeiten ist von grosser Bedeutung – sei es zum besseren Verständnis von hydrologischen Prozessen, zur Ausarbeitung von Gefahrenkarten, zur Dimensionierung von Bauwerken entlang von Fliessgewässern, zur Analyse nach Hochwasserereignissen oder für versicherungstechnische Fragestellungen. Die Abschätzung von seltenen Ereignissen verlangt nach einer Extrapolation über die Beobachtungsreihen hinaus. Dies erfordert ein geeignetes statistisches Modell, um die Wahrscheinlichkeitsverteilung dieser Zufallsvariablen abzuschätzen. Die Extremwerttheorie bietet hierfür die stochastische Basis (Chavez-Demoulin und Davison 2005).

Das Bundesamt für Umwelt (BAFU) berechnet seit 1986 statistische Kenngrössen zur Hochwasserwahrscheinlichkeit (Spreafico und Stadler 1986/1988, Spreafico und Aschwanden 1991a/b), welche seit 2004 auch im Internet veröffentlicht werden (www.hydrodaten.admin.ch). Das aktuell verwendete Standardverfahren stammt aus dem Jahr 1979 (DVWK 1979). Seit dessen Entwicklung haben sich die statistischen wie auch computertechnischen Möglichkeiten stark weiterentwickelt. Aus diesem Grund hat sich das BAFU entschieden, neu ein Verfahren

anzuwenden, das die technischen Möglichkeiten und die Datengrundlage besser ausschöpft. Die Methoden, welche in diesem Artikel vorgestellt werden, wurden durch die EPFL (Ecole Polytechnique Fédérale de Lausanne, Mathematics Institute for Analysis and Applications, Chair of Statistics) vorgeschlagen und durch das BAFU an seine Bedürfnisse angepasst.

Dieser Artikel stellt die wichtigsten Eigenschaften des neuen Verfahrens vor sowie mögliche Produkte, welche zukünftig auf der Webpage des BAFU publiziert werden sollen. Hauptziel ist es, von der Leserschaft Rückmeldungen zur Praxistauglichkeit der neuen Methodik zu erhalten.

Die Methodik wird am Fallbeispiel des Flusses Muota vorgestellt. Die verwendeten Abflussdaten wurden an der eidgenössischen hydrologischen Messstation in Ingenbohl in der Innerschweiz erhoben. Die Messstation ist seit 1917 in Betrieb, allerdings war sie bis 1923 nicht mit einem Limnigraphen ausgerüstet und der Wasserstand wurde nur einmal täglich abgelesen (BAFU 2009). Da die Extremwertstatistik Abflussspitzen, beispielsweise Tagesmaxima oder Jahresmaxima, auswertet, werden für diese Studie nur die Abflussdaten der Periode 1923–2010 verwendet. Die Station Muota-Ingenbohl wurde als Teststation ausgewählt, da

sie eine der längsten, digital verfügbaren Messreihen der Schweiz aufweist. Vorteilhaft ist ausserdem, dass der hydrologische Charakter des Einzugsgebiets und Gründe für einige Nicht-Stationaritäten bekannt sind. Die Muota weist ein Abflussregime «nival de transition» mit hohen Abflüssen im Sommer und tiefen im Winter auf (Weingartner und Aschwanden 1989). Im Jahr 1970 wurde das Wasserkraftwerk Sahli-Stufe Glattalp in Betrieb genommen, welches seither den Abfluss der Muota beeinflusst.

Bei der Mehrheit der Messstationen des BAFU stehen Jahresmaxima (bzw. Monatsmaxima) des Abflusses ab Messbeginn, Tagesmaxima aber erst ab 1974 digital zur Verfügung. Nur bei sehr wenigen Stationen, unter ihnen auch die Station Muota-Ingenbohl, wurden die analogen Aufzeichnungen (Limnigramme) aus der Zeit vor 1974 digitalisiert. Um aufzuzeigen, welchen Einfluss die fehlende Verfügbarkeit von kontinuierlichen Daten auf die Resultate hat, werden in dieser Studie einerseits die vollständige Datenreihe der Muota-Ingenbohl von 1923 bis 2010 («lange Periode») und andererseits eine gekürzte Datenreihe von 1974 bis 2010 («kurze Periode») analysiert.

2. Methodenüberblick

Die neue Methodik, welche das BAFU zukünftig in der Extremwertstatistik einsetzen will, besteht aus zwei Teilen (Tabelle 1). Der erste, semi-parametrische Teil ermöglicht eine detaillierte explorative Datenanalyse, welche Nicht-Stationaritäten aufdeckt. Diese Erkenntnisse verbessern generell das Verständnis der Daten und dienen der Einschätzung der Validität und zeitlichen Beständigkeit der HQ -Abschätzungen, welche mit den Ansätzen des zweiten Teils berechnet wurden.

Der zweite, parametrische Teil beinhaltet drei Ansätze, um die Hochwasserwahrscheinlichkeit HQ_T abzuschätzen: der Block-Maximum (BLOCK-), r -Largest-Order (R5LARG-) und der Peaks-Over

Ansatz		Stationär	Nicht-stationär	Modelleingangsdaten	Anwendung
Semi-parametrisches Verfahren	Ansatz von Chavez-Demoulin und Davison (2005), M.-O. Boldi	-	Kap. 4	Tagesmaxima	Explorative Datenanalyse
Parametrische Verfahren	Block Maximum (BLOCK)	Kap. 3.1.1	Kap. 3.2	Jahresmaxima	Abschätzung von HQ_T
	r -Largest Order Statistik (R5LARG)	Kap. 3.1.2	Kap. 3.2	Tagesmaxima	Abschätzung von HQ_T
	Peaks Over Threshold (POT)	Kap. 3.1.3	Kap. 3.2	Tagesmaxima	Abschätzung von HQ_T

Tabelle 1. Übersicht über das vom BAFU neu verwendete Methodenset.

Threshold (POT)-Ansatz. Sie unterscheiden sich in der Methode, wie die Extremwerte identifiziert werden.

Für ein besseres Verständnis werden im Folgenden zuerst die parametrischen Verfahren (Kap. 3) und aufbauend darauf das semi-parametrische Verfahren erklärt (Kap. 4).

3. Parametrische Verfahren

3.1 Stationärer Fall

3.1.1 Block-Maximum-Ansatz (BLOCK)

Der Block-Maximum-Ansatz (BLOCK) definiert den höchsten Wert innerhalb eines Zeitintervalls (Block) als Extremwert. In dieser Studie wird eine Blockgrösse von einem Jahr gewählt. Wie beim bisher verwendeten DVWK-Verfahren (1979) werden also Jahresmaxima ausgewertet. Die Extremwerttheorie zeigt, dass die Wahrscheinlichkeitsverteilung von Jahresmaxima der Allgemeinen Extremwertverteilung (Generalized Extreme Value Distribution, GEV) folgt. Die GEV wird durch den Lageparameter μ , den Skalenparameter σ und den Formparameter κ beschrieben. Durch Maximierung der log-Likelihood-Funktion, welche aus dem GEV-Modell folgt, können die drei Parameter geschätzt werden (Formel 1).

$$l(\mu, \sigma, \kappa) = -\frac{1}{\kappa} \sum_{i=1}^n \log \left(1 + \kappa \frac{y_i - \mu}{\sigma} \right) \quad (1)$$

y_i entspricht dem Jahresmaximum des Jahres i . Anhand der Schätzer für μ , σ und κ wird schliesslich mit Formel 2 der T -jährige Erwartungswert des Abflusses HQ_T ($T=1/p$) bestimmt.

$$HQ_T = \begin{cases} \mu - \frac{\sigma}{\kappa} [1 - \{-\log(1-p)\}^{-\kappa}], & \text{if } \kappa \neq 0 \\ \mu - \sigma \log\{-\log(1-p)\}, & \text{if } \kappa = 0 \end{cases} \quad (2)$$

3.1.2 r -Largest-Order-Statistik (R5LARG)

Bei der r -Largest-Order-Statistik werden die r höchsten Messwerte innerhalb eines bestimmten Zeitintervalls analysiert. Dieser Ansatz ist die Verallgemeinerung des BLOCK-Ansatzes: Bei gleicher Wahl des Zeitintervalls und $r=1$ entspricht die r -Largest-Order-Statistik der Blockmaximum-Methode. Die Grundidee ist, dass durch den Beizug von mehr Werten pro Datenblock der Informationsgehalt steigt und damit die Datengrundlage besser genutzt wird. Je grösser r ist, desto sicherer werden die Schätzungen – allerdings besteht die Gefahr eines Bias, wenn Nicht-Extremereignisse in die Analyse einfließen.

Die Wahl von r geschieht meist heuristisch. Verschiedene Tests am BAFU haben ergeben, dass bei Verwendung von Abflusstagesmaxima und einer Blockgrösse von einem Jahr mit $r=5$ ein guter Kompromiss eingegangen wird. Der Ansatz wird im Folgenden R5LARG genannt.

Die Parameter werden durch Maximierung folgender log-Likelihood-Funktion abgeschätzt (Formel 3).

$$l(\mu, \sigma, \kappa) = -\left(1 + \kappa \frac{m_r - \mu}{\sigma}\right)^{-\frac{1}{\kappa}} - r \log \sigma - \sum_{i=1}^r \left(\frac{1}{\kappa} + 1\right) \log \left(1 + \kappa \frac{m_i - \mu}{\sigma}\right) \quad (3)$$

Die Werte $m_1 \geq \dots \geq m_r$ entsprechen den r grössten Abflussereignissen pro Jahr. Da die Parameter beim R5LARG-Ansatz jenen des BLOCK-Ansatzes entsprechen, kann dieselbe Formel zur Abschätzung der HQ_T verwendet werden (Formel 2).

Die log-Likelihood-Funktion (Formel 3) darf nur unter der Annahme verwendet werden, dass die Tagesmaximadaten unabhängig voneinander sind. Daher müssen die Daten vorher von Clustern berei-

nigt werden. Beard (Cunnane 1989 zitiert in Meylan et al. 2008) berechnet den Zeitabstand L zwischen zwei unabhängigen Ereignissen über die Einzugsgebietsfläche A in Quadratmeilen (Formel 4).

$$L = 5 \text{ Tage} + \ln(A) \quad (4)$$

Die Untersuchung der BAFU-Abflussdaten hat ergeben, dass sich diese Formel gut zur Abgrenzung unabhängiger Ereignisse eignet, sofern ein Mindestabstand von 11 Tagen eingehalten wird.

3.1.3 Peaks-Over-Threshold (POT)

Beim POT-Ansatz gelten Werte als Extremereignisse, wenn sie einen hohen Schwellenwert u überschreiten. Gemäss der Extremwerttheorie ist bei einem hohen u die Anzahl Schwellenwertüberschreitungen poisson-verteilt mit der Ereignisrate λ . Die Beträge der Schwellenwertüberschreitungen folgen der Allgemeinen Pareto-verteilung (Generalized Pareto Distribution, GPD). Im Verfahren, wie es das BAFU anwendet, ist die Serie der Tagesmaxima die Eingangsgrösse. Der Schwellenwert u wird anhand der grafischen Methoden «mean residual life plot» und «fitrange plot» (vgl. Coles 2001) festgelegt. Danach werden die Schwellenwertüberschreitungen bestimmt und die Parameter durch Maximierung der log-Likelihood-Funktion, abgeleitet aus der Poisson- und GPD-Verteilung, geschätzt (Formel 5).

$$l(\lambda, \sigma, \kappa) = n \log \lambda - n \log \sigma -$$

$$\left(1 + \frac{1}{\kappa}\right) \sum_{i=1}^n \log \left(1 + \frac{\kappa w_i}{\sigma}\right) \quad (5)$$

w_i ist ein Element der Serie der n beobachteten Schwellenwertüberschreitungen. Parameter λ entspricht der Ereignisrate, das heisst der mittleren Anzahl Schwellenwertüberschreitungen während des Zeitintervalls von einem Jahr. Parameter σ entspricht auch beim POT-Ansatz dem Skalenparameter, ist aber numerisch nicht identisch mit dem σ des BLOCK- oder R5LARG-Ansatzes. Das HQ_T lässt sich durch Einsetzen der Parameterschätzer in Formel 6 berechnen.

$$HQ_T = u + \frac{\sigma}{\kappa} \left\{ \left(\frac{\lambda}{1-p} \right)^{\kappa} - 1 \right\}, T = \frac{1}{p} \quad (6)$$

Da es sehr wahrscheinlich ist, dass Schwellenwertüberschreitungen an aufeinander folgenden Tagen dasselbe Hoch-

wasserereignis repräsentieren, werden die Daten gleich wie beim R5LARG-Ansatz vorgängig von solchen Clustern bereinigt (Kap. 3.1.2).

3.2 Nicht-stationärer Fall

Im stationären Fall wird angenommen, dass alle Parameter der Verteilungsfunktion über den gesamten Untersuchungszeitraum konstant sind. Wenn die Datenanalyse oder die Stationsgeschichte jedoch eine zeitliche Veränderung vermuten lassen, können die Parameter der entsprechenden Verteilungsfunktion zeitlich abhängig modelliert werden. Die Verwendung von nicht-stationären Parametern führt dazu, dass sich auch die berechneten HQ_T über die Zeit verändern. Innerhalb des vom BAFU neu angewendeten Methodensets wird die Nicht-Stationarität der Parameter μ und σ untersucht. Beispielsweise geht die Modellvariante «mul» von einem linearen Trend des Lageparameters μ aus. Die Modellvariante «mujump» berücksichtigt eine nicht-kontinuierliche Veränderung von μ zum Zeitpunkt t_0 (Wendepunkt) und versucht damit abrupte, strukturelle Veränderungen im Einzugsgebiet, wie den Bau einer Talsperre, zu berücksichtigen. Alle Modellvarianten, welche vom BAFU beim BLOCK- und R5LARG-Ansatz untersucht werden, sind in *Tabelle 2* aufgelistet. *Tabelle 3* zeigt die Modellvarianten des POT-Ansatzes.

3.3 Modelldiagnostik

Die Güte der Modelle kann grafisch mit den sogenannten Diagnostikplots, einem Wahrscheinlichkeits- und einem Quantilplot, beurteilt werden (z.B. *Bild 4*). Je besser die Daten auf der Einheitsdiagonale liegen, desto eher kann davon ausgegangen werden, dass die Modellannahmen korrekt sind (vgl. Coles 2001).

Anhand der Devianzstatistik, bei der die log-Likelihood verschachtelter Modellvarianten verglichen wird (vgl. Coles 2001), wird überprüft ob sich die Modellvarianten signifikant unterscheiden. Ist dies nicht der Fall, wird das einfachere Modell gewählt.

Der Unsicherheit der Berechnungen wird mit Konfidenzintervallen Rechnung getragen. Das BAFU berechnet diese anhand der Deltamethode. Die Methode basiert auf einer Taylor-Entwicklung der Funktion zur Berechnung der HQ_T und einer asymptotischen Normalverteilung. Die Güte der so berechneten Intervalle ist schwierig einzuschätzen und von Fall zu Fall unterschiedlich. Die Deltamethode ist aber für alle, auch nicht-stationären

Modellvarianten	Parameter (t: Jahr)	Anzahl Parameter
Stat	μ, σ, κ	3
Mul	$\mu=a_1+a_2*t, \sigma, \kappa$	4
Muq	$\mu=a_1+a_2*t+a_3*t^2, \sigma, \kappa$	5
Sigl	$\mu, \sigma=b_1+b_2*t, \kappa$	4
Musigl	$\mu=a_1+a_2*t, \sigma=b_1+b_2*t, \kappa$	5
Mujump	$\mu=a_1+a_2*i, \sigma, \kappa; i=0 \text{ für } t \leq t_0, i=1 \text{ für } t > t_0$	4

Tabelle 2. Verwendete Modellvarianten des BLOCK- und R5LARG-Ansatzes, μ : Lage (μ), σ : Skala (σ), κ : Form (κ)).

Modellvarianten	Parameter (t: Jahr)	Anzahl Parameter
Stat	σ, κ	2
Sigl	$\sigma=b_1+b_2*t, \kappa$	3
Sigjump	$\sigma=b_1+b_2*i, \kappa; i=0 \text{ für } t \leq t_0, i=1 \text{ für } t > t_0$	3

Tabelle 3. Verwendete Modellvarianten des POT-Ansatzes, σ : Skala (σ), κ : Form (κ)).

Modellvarianten einfach verwendbar. Für technische Details verweisen wir auf Coles (2001).

4. Semi-parametrisches Verfahren

Das semi-parametrische Verfahren ist im Wesentlichen eine Erweiterung des nicht-stationären POT-Ansatzes. Anders als bei den parametrischen Verfahren, bei welchen eine Verteilung mit bestimmten Parametern vorgegeben wird, sind nicht-parametrische Verfahren verteilungsfrei. Es werden keine Hypothesen über die Veränderung der Parameter λ , σ und κ aufgestellt, sondern die Daten selber sind ausschlaggebend für die funktionale Form. Es ist jedoch unrealistisch, dass ein hypothesenfreies Modell in der Praxis anwendbar ist. Ein solches Modell wäre entweder unmöglich an die Daten anzupassen oder aber das Modell wäre überangepasst, so dass das Modell nicht mehr aussagekräftig wäre. Chavez-Demoulin und Davison (2005) haben deshalb ein semi-parametrisches Verfahren verwendet. Der Ansatz wird anhand des Parameters κ illustriert. Das Vorgehen ist für die Parameter λ und σ analog. Um den Verlauf von κ zu modellieren wird ein Verallgemeinertes Additives Modell (Generalized Additive Model, GAM) verwendet. Das Modell ist die Summe aus zwei zeitabhängigen Komponenten, wovon die erste die Saisonalität innerhalb eines Jahres und die zweite Trends von einem Jahr zum nächsten berücksichtigt (Formel 7).

$$\kappa(t) = g_1(\text{Day of Year}(t)) + g_2(\text{Year}(t)) \quad (7)$$

Day of Year ist ein Laufindex von 0 bis 1, wobei 0 für den ersten Tag im Jahr, also den 1. Januar und 1 für den letzten Tag im Jahr,

also den 31. Dezember steht. Jede Komponente ist eine Smoothing-Spline-Kurve deren Flexibilität über den Freiheitsgrad df (degree of freedom) kontrolliert wird. In der Literatur werden verschiedenste Möglichkeiten für die automatische Wahl von Smoothing-Parametern und den dazugehörenden Freiheitsgraden vorgeschlagen (vgl. Ruppert et al. 2003). Jedoch ist es hinsichtlich der Komplexität von solchen Verfahren effizienter, den df grafisch zu bestimmen. Durch verschiedene Tests an Abflussdaten des BAFU hat sich gezeigt, dass es sinnvoll ist, für alle Messstationen dieselben, fix festgelegten Freiheitsgrade zu untersuchen. Am geeignetsten erwiesen sich $df = 2$, $df = 5$ und $df = 10$. Im Resultateteil dieses Artikels (Kap. 5) wird hauptsächlich auf die Beschreibung der Resultate mit $df = 5$ eingegangen.

Die Kurven für jeden Parameter werden durch Maximierung der Likelihood-Funktion angepasst unter Verwendung eines lokalen Scoring-Algorithmus, welcher auf der log-Likelihood-Funktion des POT-Ansatzes beruht. Der Algorithmus erfordert aufwendige numerische Techniken, wobei die Annäherungsprozesse teilweise nicht robust sind. Die Konfidenzintervalle werden mit einem Bootstrap-Verfahren berechnet.

Durch Einsetzen der geschätzten Parameterkurven in die Formel zur Berechnung der HQ_T gemäss POT-Ansatz erhält man für jeden einzelnen Tag des Untersuchungszeitraums einen Wert. Diese Werte sind nicht vergleichbar mit den HQ_T aus den parametrischen Verfahren, welche für eine ganze Periode (stationärer Fall) bzw. ein einzelnes Jahr (nicht-stationärer Fall) gelten. Der semi-parametrische Ansatz ist daher auf die Datenexploration beschränkt. Anders als bei den parametrischen Verfahren ist es unmöglich, Prognosen ausserhalb der beobachteten Zeit-

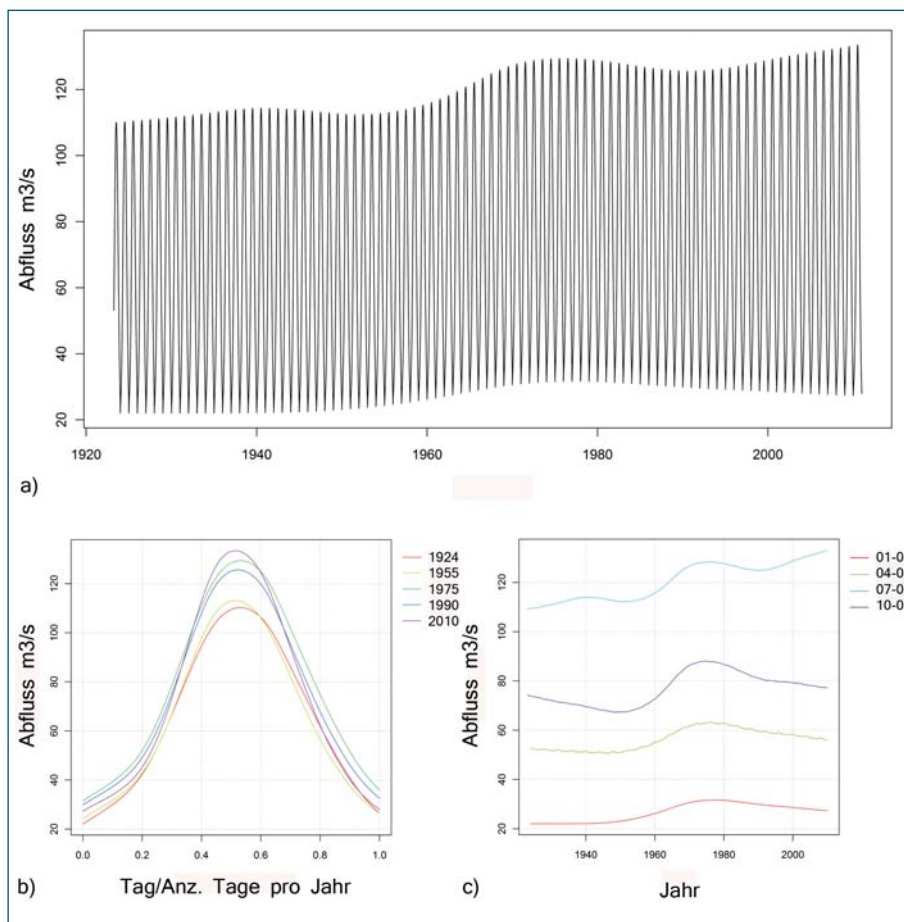


Bild 1. HQ_{30} [m³/s] für Muota-Ingenbohl, berechnet mit dem semi-parametrischen Verfahren (lange Periode, $df = 5$): a) Gesamtbetrachtung, b) in spezifischen Jahren, c) an spezifischen Tagen.

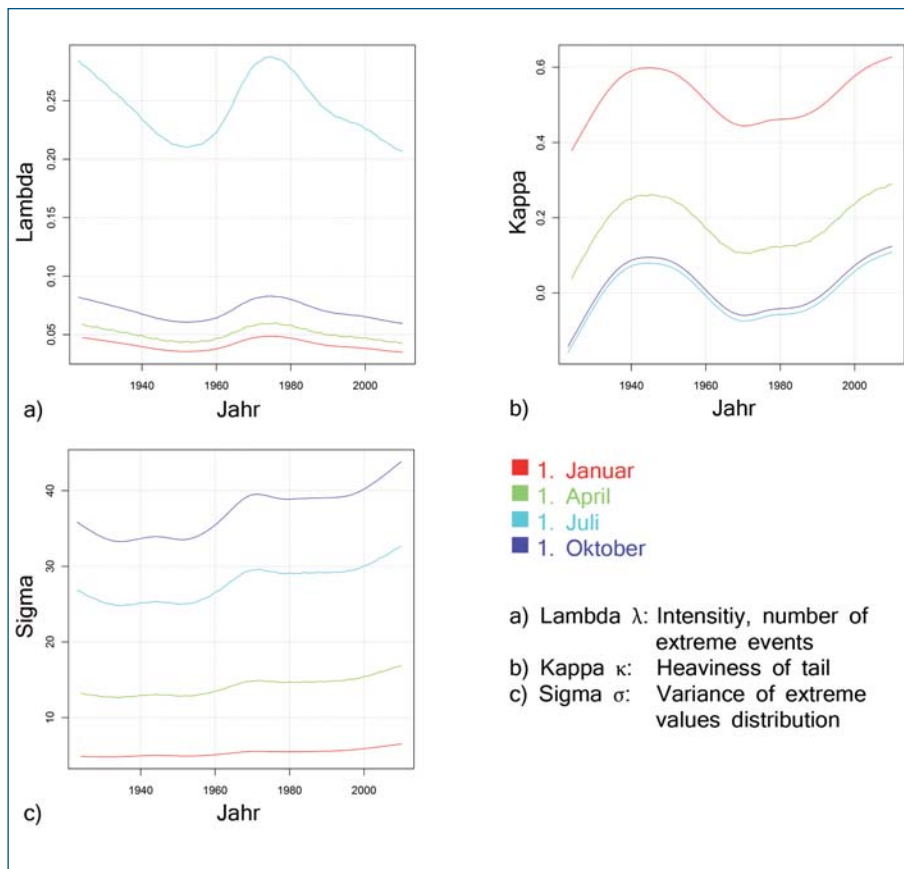


Bild 2. Parameterkurven des semi-parametrischen Verfahrens, angepasst an die Abflussdaten der Muota-Ingenbohl (lange Periode, $df = 5$).

periode zu machen. Ausserdem sind auch keine statistischen Tests beispielsweise bezüglich Trends möglich.

5. Resultate der Fallstudie Muota-Ingenbohl

5.1 Semi-parametrisches Verfahren

Resultate für die lange Periode

Die Kurven der modellierten HQ_T und λ , σ und κ können optisch in mehreren Dimensionen untersucht werden, wie dies in den Bildern 1 und 2 dargestellt ist.

Die ausgeprägte Saisonalität, mit hohen Abflüssen im Sommer und niedrigen im Winter, widerspiegelt sich deutlich im Verlauf der HQ_{30} (Bild 1a und b). Die Modellierung der Saisonalität ermöglicht, dass das HQ_{30} im Winter tiefer liegt als im Sommer; ein bestimmter Abfluss, welcher im Sommer gewöhnlich ist, also im Winter als Extremereignis erkannt wird.

Zusätzlich zeigt Bild 1, wie sich die modellierten HQ_{30} über die Jahre verändern. Während die Werte zwischen den 1920er- und den 1950er-Jahren ungefähr innerhalb derselben Bandbreite variieren, steigen sie ab den 1960er-Jahren an und pendeln sich ab Mitte der 1970er-Jahre wieder auf einem höheren Niveau ein. Eine fokussierte Betrachtung der Saisonalität in ausgewählten Jahren innerhalb des Untersuchungszeitraums (Bild 1b) zeigt, dass im Sommer die Erhöhung der modellierten HQ_{30} besonders ausgeprägt ist und dass 2010 das Maximum früher erreicht wird als in den Jahren zuvor. Der Langzeitverlauf der Werte, betrachtet an einzelnen Tagen (Bild 1c) deckt eine wesentliche Veränderung in den 1960er- bis 1980er-Jahren auf. Der Erwartungswert für ein 30-jährliches Ereignis am 1. Juli steigt seither stetig an. An den anderen untersuchten Tagen sinkt der Erwartungswert nach einem Anstieg um ca. 1975 wieder auf ein ähnliches Niveau wie zu Beginn der Untersuchungsperiode.

Ein vertieftes Verständnis der Daten erhält man bei der Analyse der Parameterkurven (Bild 2). Parameter λ beschreibt die Ereignisrate (Bild 2a). Der saisonal angepasste Schwellenwert u , der zur Ausscheidung von Extremereignissen herangezogen wird, wird generell am 1. Juli deutlich häufiger überschritten als an den anderen drei Tagen. Allerdings variiert die Ereignishäufigkeit am 1. Juli, welcher hier als stellvertretend für den Sommer betrachtet wird, über die Jahre stark. Der Formparameter κ steht direkt im Zusammenhang mit der sogenannten Hea-

viness of Tail der Verteilung. Je höher κ , desto schwerer ist der Tail, was bedeutet, dass sich der rechte Ast der Verteilungsfunktion bis zu sehr hohen Abflusswerten erstreckt. Bei negativen Werten für κ wird die Extremwertverteilung als endlich angesehen. Alle Kurven in Bild 2b zeigen einen anfänglich minimalen Anstieg von κ ab circa 1970, welcher ab ungefähr 1990 stark zunimmt. Vor 1930 war κ am 1. Oktober und am 1. Juli negativ. Parameter σ , ein Mass für die Varianz der Verteilungsfunktion, ist im Oktober am höchsten und im Januar am tiefsten (Bild 2c). Im Verlauf der Jahre hat σ an allen untersuchten Tagen zugenommen.

Der Verlauf der drei Parameter λ , κ und σ und der modellierten HQ_{30} lässt darauf schliessen, dass 2010 Hochwasserereignisse an der Muota in Ingenbohl insbesondere im Sommer seltener sind als zu Beginn der Untersuchungsperiode, die Varianz und das Ausmass der Ereignisse jedoch stark gestiegen sind. Im Bereich der 1970er-Jahre verändert sich die Abflusscharakteristik der Muota markant, was wahrscheinlich auf den Bau des KW Glattalp zurückzuführen ist.

Resultate für die kurze Periode

In der kurzen Periode können keine markanten Veränderungen der modellierten HQ_{30} über die Zeit festgestellt werden (Bild 3). 1975–2010 ist der Verlauf in der kurzen und der langen Periode ähnlich.

5.2 Parametrische Verfahren

5.2.1 Block-Maximum-Ansatz (BLOCK)

Resultate für die lange Periode

Die Diagnostikplots zeigen, dass das stationäre Modell mittlere Abflussmaxima gut abbildet, jedoch im Bereich der tiefen und insbesondere auch der sehr hohen Jahresmaxima die Anpassung an die beobachteten Werte eher schlecht gelingt. Die nicht-stationären Modellvarianten «mul» und «mujump» bringen gemäss Devianzstatistik und grafischer Beurteilung eine signifikante Verbesserung der Modellgüte. Obwohl Modellvariante «mul» eine leicht bessere log-Likelihood aufweist als «mujump», wird im Wissen über die Inbetriebnahme des KW Glattalp eher von einer abrupten als von einer kontinuierlichen Veränderung ausgegangen und deshalb «mujump» bevorzugt (Bild 4). Die mit einem nicht-stationären Modell berechneten HQ_7 gelten jeweils nur für ein bestimmtes Jahr (Kap. 3.2); dargestellt wird das aktuellste Jahr, im vorliegenden Artikel 2010. Auf-

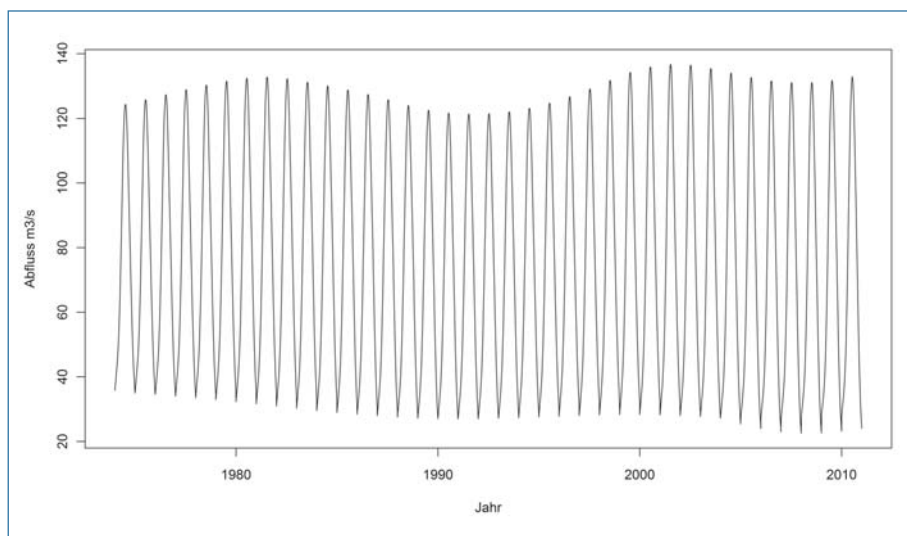


Bild 3. HQ_{30} [m³/s] für Muota-Ingenbohl, berechnet mit dem semi-parametrischen Verfahren (kurze Periode, $df = 5$).

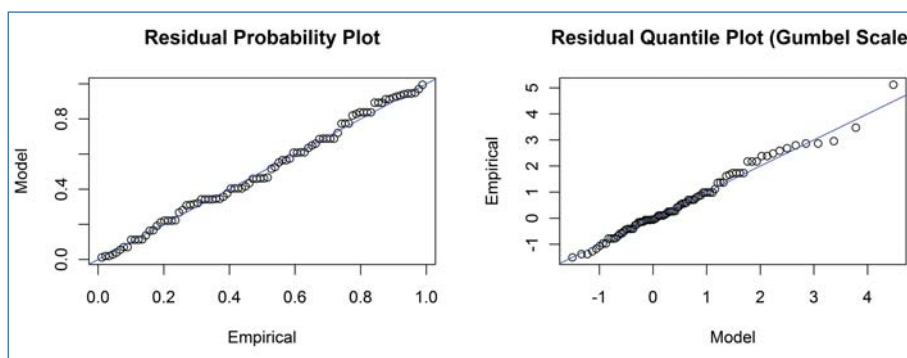


Bild 4. Diagnostikplots der Modellvariante «mujump» mit Wendepunkt 1970 des BLOCK-Ansatzes, Muota-Ingenbohl, lange Periode. Modellparameter: μ vor 1970: 141.9257, μ nach 1970: 162.1620, σ : 30.5639, κ : 0.1985.

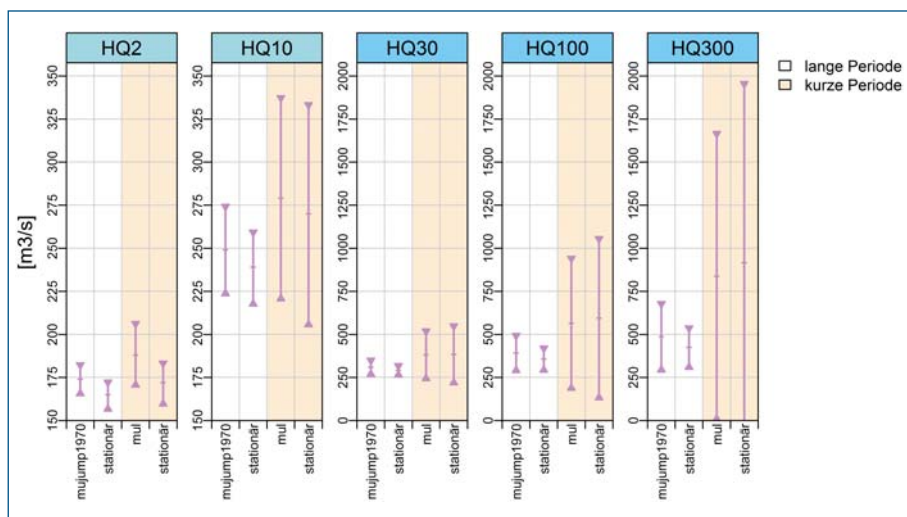


Bild 5. HQ_7 [m³/s] für Muota-Ingenbohl, berechnet mit einer nicht-stationären Modellvariante und der stationären Modellvariante des BLOCK-Ansatzes für das Jahr 2010, lange Periode (N: 88), kurze Periode (N: 37). Konfidenzintervalle gemäss Deltamethode, $\alpha = 0.05$.

grund der verbesserten Modellierung von grossen Jahresmaxima liegen die HQ_7 generell höher als beim stationären Modell, allerdings sind auch die Konfidenzintervalle grösser (Bild 5).

Resultate für die kurze Periode

Wenn nur mit der gekürzten Datenserie ab 1974 gearbeitet wird, kann der Betriebsbeginn des Kraftwerks in den Daten nicht festgestellt werden. Aufgrund der Befunde aus dem semi-parametrischen Ansatz

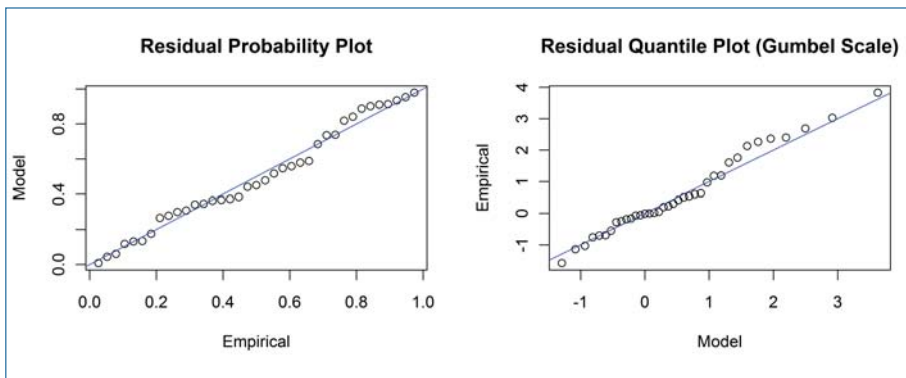


Bild 6. Diagnostikplots der Modellvariante «mul» des BLOCK-Ansatzes, Muota-Ingenbohl, kurze Periode. Modellparameter: a_1 : 145.8375, a_2 : 0.8627, σ : 26.5763, κ : 0.4315.

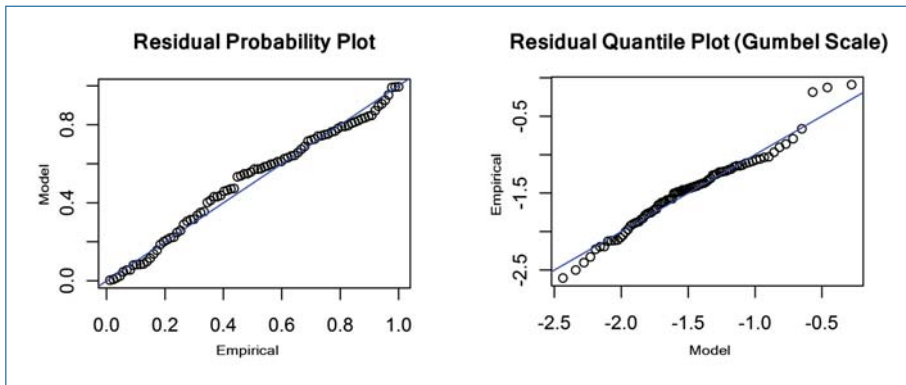


Bild 7. Diagnostikplots der Modellvariante «mujump» mit Wendepunkt 1970 des R5LARG-Ansatzes, Muota-Ingenbohl, lange Periode. Modellparameter: μ vor 1970: 149.6835, μ nach 1970: 167.0304, σ : 40.8593, κ : -0.0539.

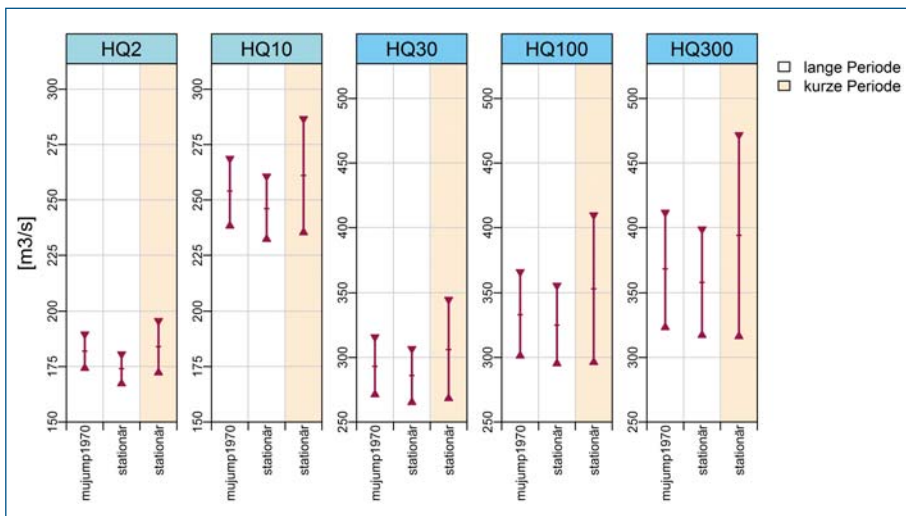


Bild 8. HQ_T [m³/s] für Muota-Ingenbohl, berechnet mit einer nicht-stationären Modellvariante und der stationären Modellvariante des R5LARG-Ansatzes für das Jahr 2010, lange Periode (N: 435), kurze Periode (N: 185). Konfidenzintervalle gemäss Deltamethode, $\alpha = 0.05$.

wurde für die Modellvariante «mujump» als Wendepunkt das Jahr 1990 ausgewählt (Bild 3). Von allen untersuchten Modellvarianten unterscheidet sich aber einzig die Variante «mul» signifikant von «stat». Wie die Diagnostikplots zeigen, ist deren Anpassung insbesondere an hohe Werte recht gut (Bild 6). Die Konfidenzintervalle für ein HQ_2 und HQ_{10} sind bei der Modell-

variante «mul» grösser als bei «stat», für ein HQ_{30} , HQ_{100} und HQ_{300} ist dies jedoch umgekehrt (Bild 5).

5.2.2 r-Largest-Order-Statistik (R5LARG)

Resultate für die lange Periode

Die Devianzstatistik deckt auf, dass sich

alle nicht-stationären Modellvarianten signifikant vom stationären Modell unterscheiden, wobei die Modellauswahl auf «mul» und «mujump» mit Wendepunkt 1970 eingeschränkt werden kann. Der Entscheid zwischen diesen Modellvarianten muss schliesslich hydrologisch begründet werden. Die Wahl fällt wie bei der BLOCK-Methode auf die Modellvariante «mujump» (Bild 7). Bild 8 zeigt, dass die HQ_T berechnet mit «mujump» generell leicht tiefer liegen als beim stationären Modell. Die Konfidenzintervalle sind grundsätzlich bei beiden Varianten sehr eng, wenn auch bei «mujump» etwas grösser.

Resultate für die kurze Periode

Die Devianzstatistik zeigt bei keiner der nicht-stationären Modellvarianten signifikante Unterschiede zum stationären Modell. Die Modellwahl fällt deshalb auf «stat». Von dieser Modellvariante wird also angenommen, dass sie die Realität am besten abbildet. Die Diagnostikplots zeigen jedoch, dass auch mit dem stationären Modell besonders im Extrembereich keine ideale Anpassung an die beobachteten Abflüsse gelingt. Die Konfidenzintervalle sind hingegen auffallend klein (Bild 8).

5.2.3 Peaks Over Threshold (POT)

Resultate für die lange Periode

Die drei untersuchten Modellvarianten sind im Bereich der tiefen und mittleren Abflüsse recht gut, die Anpassung an sehr hohe Abflüsse ist jedoch nicht ideal. Da die Devianzstatistik keinen signifikanten Unterschied aufdeckt, wird die einfachste Modellversion «stat» ausgewählt (Bild 9). Die Erwartungswerte, welche basierend auf «stat» berechnet wurden, sind aus Bild 10 ersichtlich.

Resultate für die kurze Periode

Werden nur Daten ab 1974 verwendet, unterscheiden sich die drei Modellvarianten wie in der langen Periode nicht signifikant voneinander. Die Wahl fällt auf das einfachste Modell «stat».

5.2.4 Zusammengefasste Resultate der parametrischen Verfahren

Bild 11 zeigt die HQ_T jener Modellvarianten, welche jeweils am besten an die Daten angepasst waren. Sowohl die Werte wie auch die Grösse der Konfidenzintervalle, welche ein Mass für die Unsicherheit sind, variieren stark. Generell werden diese Unterschiede grösser, je höher die Wiederkehrperiode T ist. Der R5LARG-Ansatz zeigt bei allen Wiederkehrperioden durch-

gänglich die kleinsten Konfidenzintervalle, was vermuten lässt, dass dieser Ansatz für die Daten der Muota vergleichsweise robust ist. Die HQ_2 und HQ_{10} der einzelnen Ansätze streuen um weniger als $10 \text{ m}^3/\text{s}$. Die Differenz zwischen einem HQ_{100} berechnet mit R5LARG bzw. BLOCK beträgt jedoch beinahe $60 \text{ m}^3/\text{s}$. Die Konsequenz ist, dass ein HQ_{100} gemäss BLOCK-Ansatz höher liegt als das HQ_{300} des R5LARG-Ansatzes. Die HQ_7 für die kurze Periode liegen, unabhängig vom gewählten Ansatz, durchwegs höher als jene für die lange Periode und die Konfidenzintervalle sind, aufgrund der verringerten Datenbasis, grösser (Bilder 5, 8 und 10).

6. Schlussfolgerungen und Ausblick

Im Vergleich mit dem bisher vom BAFU verwendeten Standardverfahren (DVWK 1979) verbessert das neue Set von Auswertungsmethoden die Ausschöpfung der Datengrundlage, ermöglicht eine vertiefte Datenanalyse und die Berücksichtigung von Nicht-Stationaritäten.

Das semi-parametrische Verfahren erlaubt interessante Aussagen über Veränderungen der Häufigkeit, des Ausmasses und der Variabilität von Hochwasserereignissen. Dies gilt insbesondere für lange, rückdigitalisierte Reihen. Die berechneten HQ_7 sind aber nicht direkt verwendbar, da sie nur für einen expliziten Tag in einem expliziten Jahr gelten.

Im Fallbeispiel der Muota kann mit nicht-stationären Varianten der parametrischen Ansätze BLOCK und R5LARG die Modellgüte verbessert werden, da die Zunahme extremer Hochwasser in den letzten Jahren einbezogen wird. Die einfachen mathematischen Formeln zur Berücksichtigung von Nicht-Stationaritäten (Tabelle 2 und 3) bilden die Realität allerdings nur annäherungsweise ab. Zudem bleibt unklar, ob sich ein Trend in der Zukunft unverändert fortsetzt. Die berechneten HQ_7 gelten jeweils nur für einen bestimmten Zeitpunkt, in diesem Artikel für das Jahr 2010.

Der grundlegende Unterschied zwischen allen parametrischen Verfahren liegt bei der Extremwertidentifikation. Der BLOCK-Ansatz verwendet grundsätzlich fünfmal weniger Werte als R5LARG. Da bei den meisten eidgenössischen Abflussmessstationen die Grundlagendaten für R5LARG erst ab 1974, für BLOCK dagegen meist wesentlich weiter zurück vorliegen, ist der Unterschied im Stichprobenumfang N in der Praxis häufig etwas kleiner. Bei R5LARG besteht zudem verstärkt die Gefahr, Abflussdaten auszuwer-

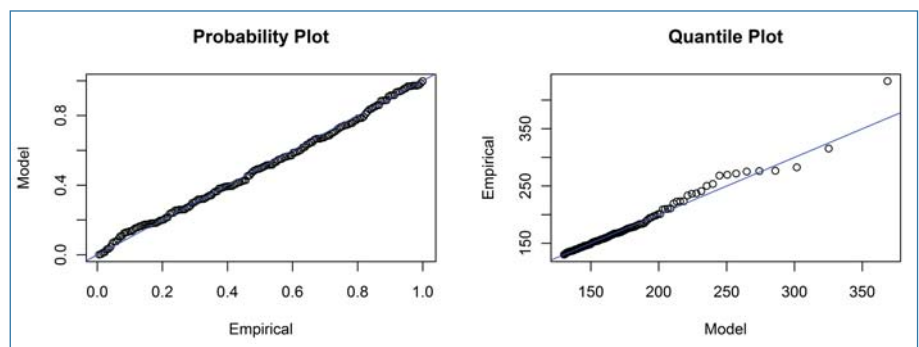


Bild 9. Diagnostikplots der Modellvariante «stat» des POT-Ansatzes, Muota-Ingenbohl, lange Periode. Modellparameter: $\sigma: 31.18$, $\kappa: 0.14$.

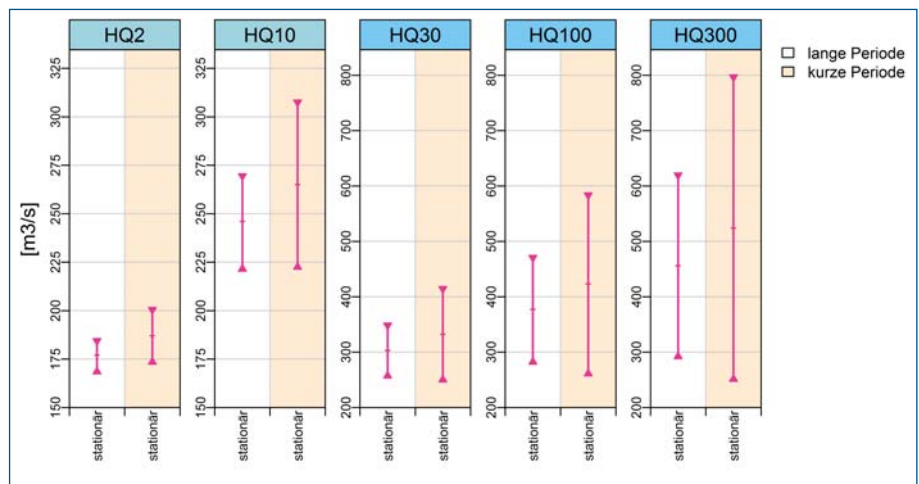


Bild 10. $HQ_T [\text{m}^3/\text{s}]$ für Muota-Ingenbohl, berechnet mit der stationären Modellvariante des POT-Ansatzes, lange Periode (N: 171), kurze Periode (N: 120). Konfidenzintervalle gemäss Deltamethode, $\alpha = 0.05$.

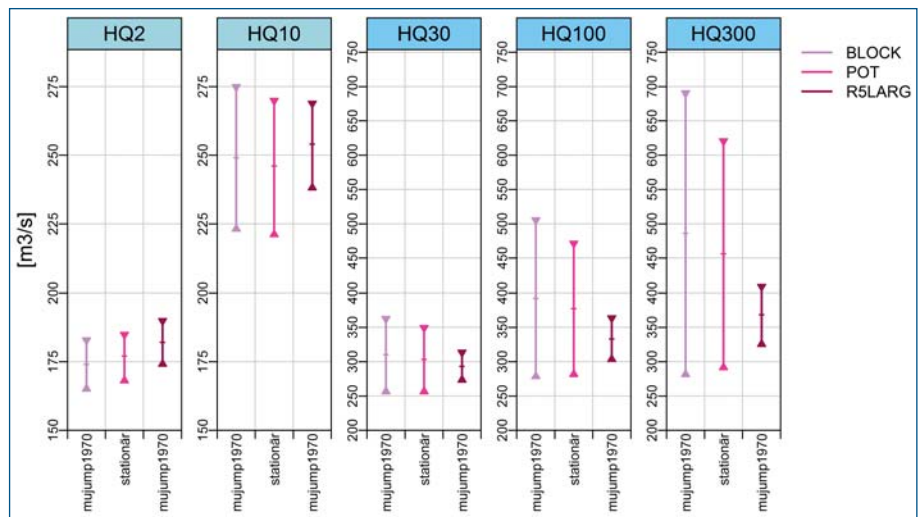


Bild 11. $HQ_T [\text{m}^3/\text{s}]$ für Muota-Ingenbohl, berechnet mit den jeweils bestangepassten Modellvarianten des BLOCK-, POT- bzw. R5LARG-Ansatzes, lange Periode.

ten, die nicht extrem sind. Dieses Risiko kann beim POT-Ansatz mit einer guten Wahl des Schwellenwerts u minimiert werden. Zusätzlich wird beim POT-Ansatz die Bildung von Datenblöcken umgangen, was verhindert, dass Extremwerte weglassen werden, nur weil im selben Block ein bzw. fünf höhere Werte vorkommen (Coles 2001). Die meist fehlenden digitalen Grundlagendaten vor 1974 schränken

jedoch auch die Anwendung des POT-Ansatzes ein.

Alle Modellansätze haben Vor- und Nachteile, was eine generelle Modellwahl ausschliesst. Auch bei einer einzelnen Auswertung bleibt unbekannt, welcher Ansatz die Realität am besten widerspiegelt. Das BAFU strebt daher an, in naher Zukunft neben den HQ_7 des stationären BLOCK-Ansatzes, welcher das DVWK-

Verfahren von 1979 ablöst, weitere Resultate zu publizieren.

Die Ergebnisse der hier vorgestellten Methoden sollen in kurzen Berichten für alle aktuellen Abflussmessstationen zusammengestellt werden. Diese Unterlagen könnten in der Praxis dazu dienen, Nicht-Stationaritäten in den Daten zu erkennen und die Unterschiede in den über die verschiedenen Ansätze berechneten HQ_T zu beurteilen. Für die Beantwortung konkreter Fragestellungen, z.B. bei einer Bemessung, bilden die Resultate dieser statistischen Auswertungen nur eine Grundlage. Kenntnisse über die hydrologischen Prozesse im Einzugsgebiet, Überlegungen zu Worst-Case-Szenarien für Niederschlag und Abfluss usw. sind unerlässlich, um das Hochwasserrisiko abzuschätzen. Die Abteilung Hydrologie des BAFU befürwortet eine «hydrologisch-argumentative Vorgehensweise», wie sie Merz und Blöschl (2008 zitiert in DWA 2012:52) vorschlagen.

Zurzeit wird das vorgestellte Methodenset an Daten anderer Abflussmessstationen intensiv getestet und der Erfahrungsschatz wird erweitert. Anregungen und Rückmeldungen zum Methodenset und zu den vorgesehenen Produkten sind dem BAFU sehr willkommen.

Literatur

- BAFU (2009). Hydrologisches Jahrbuch der Schweiz 2008. Bern.
- Chavez-Demoulin, V., Davison, A. (2005). Generalized additive models for sample extremes. *Applied Statistics* 54 (1): 207–222.
- Coles, S. (2001). An Introduction to Statistical Modeling of Extreme Values. Springer. London.
- DVWK (Hrsg.) (1979). DVWK-Regel 101. Empfehlungen zur Berechnung von Hochwasserwahrscheinlichkeiten. 2. bearb. Auflage. Deutscher Verband für Wasserwirtschaft und Kulturbau e. V. Bonn.
- DWA (Hrsg.) (2012). Merkblatt DWA-M 552. Ermittlung von Hochwasserwahrscheinlichkeiten. Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V. Hennef.
- Merz, R., Blöschl, G. (2008). Flood frequency hydrology: 2. Combining data evidence. *Water Resources Research*, 44 (8), zitiert in DWA (2012).
- Meylan, P., Favre, A.-C., Musy, A. (2008). Hydrologie fréquentielle. Une science prédictive. Presses polytechniques et universitaires romandes. Lausanne.
- Ruppert, D., Wand, M. P., Carroll, R. J. (2003). Semiparametric Regression. Cambridge University Press. Cambridge.
- Spreafico, M., Stadler, K. (1986). Hochwasserabflüsse in schweizerischen Gewässern. Mitteilung der Landeshydrologie und -geologie Nr. 7. Bern.

- Spreafico, M., Stadler, K. (1988). Hochwasserabflüsse in schweizerischen Gewässern. Mitteilung der Landeshydrologie und -geologie Nr. 8. Bern.
- Spreafico, M., Aschwanden, H. (1991a). Hochwasserabflüsse in schweizerischen Gewässern. Mitteilung der Landeshydrologie und -geologie Nr. 16. Bern.
- Spreafico, M., Aschwanden, H. (1991b). Hochwasserabflüsse in schweizerischen Gewässern. Mitteilung der Landeshydrologie und -geologie Nr. 17. Bern.
- Weingartner, R., Aschwanden, H. (1989). HADES Tafel 5.2. Abflussregimes als Grundlage zur Abschätzung von Mittelwerten des Abflusses. Bundesamt für Umwelt (Hrsg.): Hydrologischer Atlas der Schweiz. Bern.

Anschrift der Verfasser

Caroline Kan, Eva Baumgartner, Simon Schick
Bundesamt für Umwelt BAFU
Abteilung Hydrologie, CH-3003 Bern
caroline.kan@bafu.admin.ch
Marc-Olivier Boldi
Ecole Polytechnique Fédérale de Lausanne,
Mathematics Institute for Analysis and Applications,
Chair of Statistics, CH-1015 Lausanne,
und University of Geneva, Faculty of Economics and Social Sciences, Research Center for Statistics, CH-1211 Geneva 4.

**That's
theWey.**



Schweizer Fachmesse für Öffentliche Betriebe + Verwaltungen

Bern, 18.–21.6.2013

Besuchen Sie uns in Halle 302, Stand D044.



Wey Plattenschieber



Wey Absperr- und
Rückschlagklappen



Wey Kanal-
absperrore

Unser breites und flexibles Programm für Kanalabsperungen, Hochwasserschutz und für die Wassertechnik kann auf individuellste Bedürfnisse angepasst werden. Technisch raffinierte Lösungen für die Wasser- und Abwasseraufbereitung, für Auffang- und Rückhaltebecken sowie für Kanal- und Schachtbauwerke machen uns auch in diesem Bereich zu einem der weltweit führenden Anbieter.

**Das Wasser im Griff. Wassertechnik
und Hochwasserschutz.**

WEY®
VALVE INNOVATION MANAGEMENT



SWISS
PERFORMANCE

SISTAG AG, Alte Kantonsstrasse 7, 6274 Eschenbach, Switzerland, Tel +41 41 449 9944, weyvalve.ch